RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/903,180_

Source: _OIPE_

Date Processed by STIC: _7/26/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: _09/903,18''_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 _____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 _____ Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 _____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. **Please ensure your subsequent submission is saved in ASCII text.**

5 _____ Variable Length

Sequence(s) _____ contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 _____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

7 _____ Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8 _____ Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 _____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 _____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence

11 _____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 _____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 _____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

RAW SEQUENCE LISTING   DATE: 07/26/2001
PATENT APPLICATION: US/09/903,180   TIME: 15:13:29

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\07262001\I903180.raw**

SEQUENCE LISTING

Does Not Comply
Correct Diskette Number

```
  5 (1) GENERAL INFORMATION:
  7       (i) APPLICANT: De Robertis, Edward M.
  8                      Bouwmeester, Tewis
 10      (ii) TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 11                      Factors
 13     (iii) NUMBER OF SEQUENCES: 10
 15      (iv) CORRESPONDENCE ADDRESS:
 16           (A) ADDRESSEE: Majestic, Parsons, Siebert & Hsue
 17           (B) STREET: Four Embarcadero Center, Suite 1100
 18           (C) CITY: San Francisco
 19           (D) STATE: California
 20           (E) COUNTRY: U.S.A.
 21           (F) ZIP: 94111-4106
 23       (v) COMPUTER READABLE FORM:
 24           (A) MEDIUM TYPE: Floppy disk
 25           (B) COMPUTER: IBM PC compatible
 26           (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 27           (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
 29      (vi) CURRENT APPLICATION DATA:
 30           (A) APPLICATION NUMBER: US/09/903,180
 31           (B) FILING DATE: 11-Jul-2001
 32           (C) CLASSIFICATION:
 34     (vii) PRIOR APPLICATION DATA:
 35           (A) APPLICATION NUMBER: US 60/020,150
 36           (B) FILING DATE: 20-JUN-1996
 38    (viii) ATTORNEY/AGENT INFORMATION:
 39           (A) NAME: Siebert, J. Suzanne
 40           (B) REGISTRATION NUMBER: 28,758
 41           (C) REFERENCE/DOCKET NUMBER: 3100.002US1
 43      (ix) TELECOMMUNICATION INFORMATION:
 44           (A) TELEPHONE: 415/248-5500
 45           (B) TELEFAX: 415/362-5418
```

ERRORED SEQUENCES

```
330 (2) INFORMATION FOR SEQ ID NO: 5:
332       (i) SEQUENCE CHARACTERISTICS:                    707 (p. 3)
333           (A) LENGTH: 979 amino acids
334           (B) TYPE: amino acid
335           (D) TOPOLOGY: linear
337      (ii) MOLECULE TYPE: peptide
339      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
341 Met Leu Leu Leu Phe Arg Ala Ile Pro Met Leu Leu Leu Gly Leu Met
342 1               5                   10                  15
344 Val Leu Gln Thr Asp Cys Glu Ile Ala Gln Tyr Tyr Ile Asp Glu Glu
```

RAW SEQUENCE LISTING   DATE: 07/26/2001
PATENT APPLICATION: **US/09/903,180**   TIME: 15:13:30

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\07262001\I903180.raw**

```
345             20              25              30
347 Glu Pro Pro Gly Thr Val Ile Ala Val Leu Ser Gln His Ser Ile Phe
348         35              40              45
350 Asn Thr Thr Asp Ile Pro Ala Thr Asn Phe Arg Leu Met Lys Gln Phe
351     50              55              60
353 Asn Asn Ser Leu Ile Gly Val Arg Glu Ser Asp Gly Gln Leu Ser Ile
354 65              70              75              80
356 Met Glu Arg Ile Asp Arg Glu Gln Ile Cys Arg Gln Ser Leu His Cys
357             85              90              95
359 Asn Leu Ala Leu Asp Val Val Ser Phe Ser Lys Gly His Phe Lys Leu
360         100             105             110
362 Leu Asn Val Lys Val Glu Val Arg Asp Ile Asn Asp His Ser Pro His
363         115             120             125
365 Phe Pro Ser Glu Ile Met His Val Glu Val Ser Glu Ser Ser Ser Val
366     130             135             140
368 Gly Thr Arg Ile Pro Leu Glu Ile Ala Ile Asp Glu Asp Val Gly Ser
369 145             150             155             160
371 Asn Ser Ile Gln Asn Phe Gln Ile Ser Asn Asn Ser His Phe Ser Ile
372             165             170             175
374 Asp Val Leu Thr Arg Ala Asp Gly Val Lys Tyr Ala Asp Leu Val Leu
375             180             185             190
377 Met Arg Glu Leu Asp Arg Glu Ile Gln Pro Thr Tyr Ile Met Glu Leu
378         195             200             205
380 Leu Ala Met Asp Gly Gly Val Pro Ser Leu Ser Gly Thr Ala Val Val
381     210             215             220
383 Asn Ile Arg Val Leu Asp Phe Asn Asp Asn Ser Pro Val Phe Glu Arg
384 225             230             235             240
386 Ser Thr Ile Ala Val Asp Leu Val Glu Asp Ala Pro Leu Gly Tyr Leu
387             245             250             255
389 Leu Leu Glu Leu His Ala Thr Asp Asp Asp Glu Gly Val Asn Gly Glu
390             260             265             270
392 Ile Val Tyr Gly Phe Ser Thr Leu Ala Ser Gln Glu Val Arg Gln Leu
393         275             280             285
395 Phe Lys Ile Asn Ser Arg Thr Gly Ser Val Thr Leu Glu Gly Gln Val
396     290             295             300
398 Asp Phe Glu Thr Lys Gln Thr Tyr Glu Phe Glu Val Gln Ala Gln Asp
399 305             310             315             320
401 Leu Gly Pro Asn Pro Leu Thr Ala Thr Cys Lys Val Thr Val His Ile
402             325             330             335
404 Leu Asp Val Asn Asp Asn Thr Pro Ala Ile Thr Ile Thr Pro Leu Thr
405             340             345             350
407 Thr Val Asn Ala Gly Val Ala Tyr Ile Pro Glu Thr Ala Thr Lys Glu
408             355             360             365
410 Asn Phe Ile Ala Leu Ile Ser Thr Thr Asp Arg Ala Ser Gly Ser Asn
411     370             375             380
413 Gly Gln Val Arg Cys Thr Leu Tyr Gly His Glu His Phe Lys Leu Gln
414 385             390             395             400
416 Gln Ala Tyr Glu Asp Ser Tyr Met Ile Val Thr Thr Ser Thr Leu Asp
417             405             410             415
```

RAW SEQUENCE LISTING   DATE: 07/26/2001
PATENT APPLICATION: US/09/903,180   TIME: 15:13:30

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\07262001\I903180.raw**

```
419 Arg Glu Asn Ile Ala Ala Tyr Ser Leu Thr Val Val Ala Glu Asp Leu
420          420             425              430
422 Gly Phe Pro Ser Leu Lys Thr Lys Lys Tyr Tyr Thr Val Lys Val Ser
423          435             440              445
425 Asp Glu Asn Asp Asn Ala Pro Val Phe Ser Lys Pro Gln Tyr Glu Ala
426     450             455             460
428 Ser Ile Leu Glu Asn Asn Ala Pro Gly Ser Tyr Ile Thr Thr Val Ile
429 465             470             475              480
431 Ala Arg Asp Ser Asp Ser Asp Gln Asn Gly Lys Val Asn Tyr Arg Leu
432              485             490              495
434 Val Asp Ala Lys Val Met Gly Gln Ser Leu Thr Thr Phe Val Ser Leu
435          500             505              510
437 Asp Ala Asp Ser Gly Val Leu Arg Ala Val Arg Ser Leu Asp Tyr Glu
438          515             520              525
440 Lys Leu Lys Gln Leu Asp Phe Glu Ile Glu Ala Ala Asp Asn Gly Ile
441     530             535             540
443 Pro Gln Leu Ser Thr Arg Val Gln Leu Asn Leu Arg Ile Val Asp Gln
444 545             550             555              560
446 Asn Asp Asn Cys Pro Val Ile Thr Asn Pro Leu Leu Asn Asn Gly Ser
447              565             570             575
449 Gly Glu Val Leu Leu Pro Ile Ser Ala Pro Gln Asn Tyr Leu Val Phe
450          580             585             590
452 Gln Leu Lys Ala Glu Asp Ser Asp Glu Gly His Asn Ser Gln Leu Phe
453          595             600             605
455 Tyr Thr Ile Leu Arg Asp Pro Ser Arg Leu Phe Ala Ile Asn Lys Glu
456     610             615             620
458 Ser Gly Glu Val Phe Leu Lys Lys Gln Leu Asn Ser Asp His Ser Glu
459 625             630             635              640
461 Asp Leu Ser Ile Val Val Ala Val Tyr Asp Leu Gly Arg Pro Ser Leu
462              645             650             655
464 Val Ala Asn Arg Met His Ala Glu Tyr Glu Arg Asp Leu Val Asn Arg
```
E--> 465   ~~930~~ *660*   ~~935~~ *665*   ~~940~~ *670*
```
467 Ser Ala Thr Leu Ser Pro Gln Arg Ser Ser Ser Arg Tyr Gln Glu Phe
```
E--> 468 ~~945~~ *675*   ~~950~~ *680*   ~~955~~ *685*   ~~960~~
```
470 Asn Tyr Ser Pro Gln Ile Ser Arg Gln Leu His Pro Ser Glu Ile Ala
```
E--> 471   *690*   ~~965~~ *695*   ~~970~~ *700*   ~~975~~
E--> 473 <u>Thr Thr Phe</u> *705*
```
780 (2) INFORMATION FOR SEQ ID NO: 9:
782      (i) SEQUENCE CHARACTERISTICS:
783          (A) LENGTH: 325 amino acids
784          (B) TYPE: amino acid
785          (D) TOPOLOGY: linear
787      (ii) MOLECULE TYPE: peptide
789      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
791 Met Val Cys Gly Ser Pro Gly Gly Met Leu Leu Leu Arg Ala Gly Leu
```
E--> 792 1   5   ~~10~~ *10*   15
```
794 Leu Ala Leu Ala Ala Leu Cys Leu Leu Arg Val Pro Gly Ala Arg Ala
795          20              25              30
797 Ala Ala Cys Glu Pro Val Arg Ile Pro Leu Cys Lys Ser Leu Pro Trp
```

(← misaligned numbers — see item 3 on Error summary sheet)

```
798          35                    40                    45
800 Asn Met Thr Lys Met Pro Asn His Leu His His Ser Thr Gln Ala Asn
801     50                  55                  60
803 Ala Ile Leu Ala Ile Glu Gln Phe Glu Gly Leu Leu Gly Thr His Cys
804 65                  70                  75                  80
806 Ser Pro Asp Leu Leu Phe Phe Leu Cys Ala Met Tyr Ala Pro Ile Cys
807                 85                  90                  95
809 Thr Ile Asp Phe Gln His Glu Pro Ile Lys Pro Cys Lys Ser Val Cys
810                 100                 105                 110
812 Glu Arg Ala Arg Gln Gly Cys Glu Pro Ile Leu Ile Lys Tyr Arg His
813         115                 120                 125
815 Ser Trp Pro Glu Asn Leu Ala Cys Glu Glu Leu Pro Val Tyr Asp Arg
816     130                 135                 140
818 Gly Val Cys Ile Ser Pro Glu Ala Ile Val Thr Ala Asp Gly Ala Asp
819 145                 150                 155                 160
821 Phe Pro Met Asp Ser Ser Asn Gly Asn Cys Arg Gly Ala Ser Ser Glu
822                 165                 170                 175
824 Arg Cys Lys Cys Lys Pro Ile Arg Ala Thr Gln Lys Thr Tyr Phe Arg
825                 180                 185                 190
827 Asn Asn Tyr Asn Tyr Val Ile Arg Ala Lys Val Lys Glu Ile Lys Thr
828         195                 200                 205
830 Lys Cys His Asp Val Thr Ala Val Val Glu Val Lys Glu Ile Leu Lys
831     210                 215                 220
833 Ser Ser Leu Val Asn Ile Pro Arg Asp Thr Val Asn Leu Tyr Thr Ser
834 225                 230                 235                 240
836 Ser Gly Cys Leu Cys Pro Pro Leu Asn Val Asn Glu Glu Tyr Ile Ile
837                 245                 250                 255
839 Met Gly Tyr Glu Asp Glu Glu Arg Ser Arg Leu Leu Leu Val Glu Gly
840                 260                 265                 270
842 Ser Ile Ala Glu Lys Trp Lys Asp Arg Leu Gly Lys Lys Val Lys Arg
843         275                 280                 285
845 Trp Asp Met Lys Leu Arg His Leu Gly Leu Ser Lys Ser Asp Ser Ser
846     290                 295                 300
848 Asn Ser Asp Ser Thr Gln Ser Gln Lys Ser Gly Arg Asn Ser Asn Pro
849 305                 310                 315                 320
851 Arg Gln Ala Arg Asn
852                 325
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/903,180

DATE: 07/26/2001
TIME: 15:13:31

Input Set : **A:\seqlist.txt**
Output Set: **N:\CRF3\07262001\I903180.raw**

```
L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:465 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:5
M:332 Repeated in SeqNo=5
L:473 M:203 E: No. of Seq. differs, LENGTH:Input:979 Found:707 SEQ:5
L:792 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:9
```